

PATENT ABSTRACTS OF JAPAN

(11)Publication number : 05-334006
(43)Date of publication of application : 17.12.1993

(51)Int.Cl.

G06F 3/06

(21) Application number : 04-144622

(71)Applicant : MITSUBISHI ELECTRIC CORP

(22) Date of filing : 04.06.1992

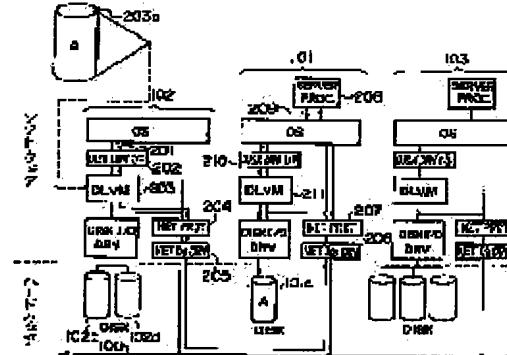
(72)Inventor : KAMIMURA JIYOZE
SAKAKURA TAKASHI

(54) LOGICAL VOLUME SYSTEM

(57)Abstract:

PURPOSE: To effectively use an idle disk space on a network by each computer device.

CONSTITUTION: With respect to the logical volume which manages physical disks 101c, 102c, 102d, etc., provided in computer devices 101 to 103 connected through a network bus 100 as logical disks, the object of a physical volume in the logical volume is extended from physical disks in the same computer system to, for example, physical disks of another system through the network bus 100.



LEGAL STATUS

[Date of request for examination] 27.03.1996

[Date of sending the examiner's decision of rejection]

[Kind of final disposal of application other than
the examiner's decision of rejection or
application converted registration]

[Date of final disposal for application]

[Patent number] 2766424

[Date of registration]

「Number of appeal against examiner's decision

[Date of requesting appeal against examiner's
of rejection]

[Decision of rejection]
[Date of the refusal]

(19)日本国特許庁 (JP)

(12) 公開特許公報 (A)

(11)特許出願公開番号

特開平5-334006

(43)公開日 平成5年(1993)12月17日

(51)Int.Cl.⁵

G 0 6 F 3/06

識別記号 庁内整理番号

3 0 1 K 7165-5B

F I

技術表示箇所

(21)出願番号

特願平4-144622

(22)出願日

平成4年(1992)6月4日

(71)出願人 000006013

三菱電機株式会社

東京都千代田区丸の内二丁目2番3号

(72)発明者 上村 ジョゼ

神奈川県鎌倉市大船五丁目1番1号 三菱
電機株式会社情報電子研究所内

(72)発明者 坂倉 隆史

神奈川県鎌倉市大船五丁目1番1号 三菱
電機株式会社情報電子研究所内

(74)代理人 弁理士 金山 敏彦 (外2名)

審査請求 未請求 請求項の数2(全8頁)

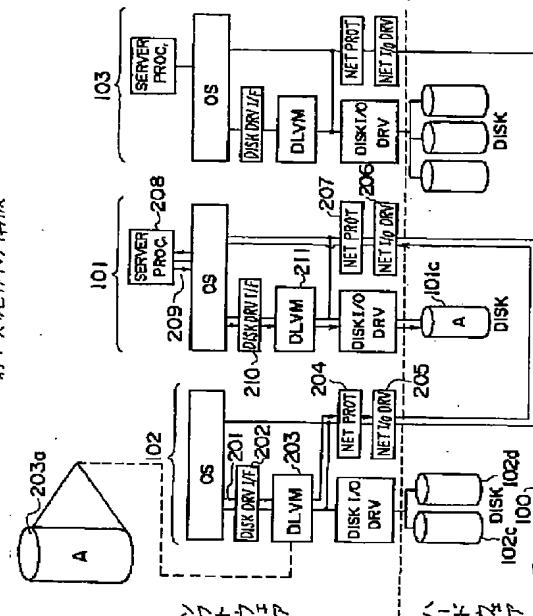
(54)【発明の名称】 論理ボリュームシステム

(57)【要約】

【目的】ネットワーク上における空きディスク空間を各コンピュータ装置で有効活用する。

【構成】ネットワークバス100を介して接続されるコンピュータ装置101～103に設けられた各物理ディスク101c, 102c, 102d等を論理ディスクとして管理する論理ボリュームにおいて、論理ボリュームにおける物理ボリュームの対象を、同一のコンピュータシステム内の物理ディスクからネットワークバス100を介して例えば他のシステムの物理ディスクに拡張するようにした。

第1実施例の構成



I

2

【特許請求の範囲】

【請求項1】 オペレーティングシステム部と、少なくとも1つのディスク装置と、そのディスク装置を管理するディスクI/O管理部と、ネットワークバスに接続されたネットワークI/O管理部と、を含むコンピュータ装置が、前記ネットワークバスに複数接続されたシステムにおいて、

少なくとも1つのコンピュータ装置において、前記オペレーティングシステム部と前記ディスクI/O管理部との間に、システム内にある1又は複数の任意のディスク装置の記憶空間を、仮想的な記憶空間である論理ボリュームとして統合管理する論理ボリューム管理部が設けられ、

前記オペレーティングシステム部は、前記論理ボリュームをディスクI/O管理部と同等とみなしてディスク制御の命令を出し、

前記論理ボリューム管理部は、前記オペレーティングシステム部からのディスク制御の命令に対応する現実のディスク装置を判断して当該ディスク装置の制御を行うことを特徴とする論理ボリュームシステム。

【請求項2】 請求項1記載の論理ボリュームシステムにおいて、

前記論理ボリューム管理部と前記ネットワークI/O管理部とが直接接続されたことを特徴とする論理ボリュームシステム。

【発明の詳細な説明】

【0001】

【産業上の利用分野】 本発明はシステム、特に、現実のディスク装置の記憶空間を仮想的な論理ボリュームとして管理する論理ボリュームシステムに関する。

【0002】

【従来の技術】 複数のコンピュータ装置がネットワークバスに接続されたシステムが各種の分野で活用されている。

【0003】 このようなシステムにおいて、各コンピュータ装置は、一般的に1又は複数のディスク装置を備えている。しかし、それぞれのコンピュータ装置において、各ディスク装置における記憶空間の使用率はばらつきがあるのが通常であり、記憶空間が不足するコンピュータ装置もあれば、記憶空間が余剰しているコンピュータ装置もある。

【0004】 したがって、コンピュータ相互がネットワーク化されている利点を生かし、デバイス資源の有効活用を図るため、各ディスク装置を複数のコンピュータで共用することが考えられ、従来のシステムでは、それが実現化されているものがある。

【0005】 ところで、複数のディスク装置の制御を簡易にするためのものとして、従来からいわゆる「論理ボリューム」という仮想記憶空間が知られている。この論理ボリュームは、通常、現実の複数のディスク装置（物

10

20

30

40

50

理ボリューム）の記憶空間を合わせた仮想的記憶空間であり、仮想上の大きなディスク装置に相当するものである。すなわち、コンピュータの基本システムからは、論理ボリュームだけがみえるようにし、その基本システムが個々のディスク装置を直接制御する煩雑さを解消するものである。

【0006】 従来、このような論理ボリュームの適応例としては、高速化のためにディスクの分割使用を行うディスクストライピング手法、より広い連続領域を得るためにディスクの並列使用を行うディスクアレイ手法あるいはディスク障害によるファイル破損に対応するためにディスクを多重化するミラーディスク手法等が提案されている。

【0007】 また、論理ボリュームに関し、特開昭63-10250号公報（文献1）や特開平3-92942号公報（文献2）等に開示されるものが挙げられる。ここで、文献1ではファイル作成時に、該当する論理ボリュームを構成する物理ボリュームを略均等に選択するようし、ファイルを各物理ボリュームに分散することにより、論理ボリュームに対応する物理ボリュームのアクセス分散を行い、アクセス効率の向上を図っている。また、文献2では、ファイルをディスク内に任意の論理的な単位で格納、多重化するようし、ディスクの利用効率の向上を図っている。

【0008】

【発明が解決しようとする課題】 しかしながら、従来においては、マルチコンピュータ環境上で、ネットワークを介した論理ボリュームは構築されていなかった。このため、システムにおけるディスクの空き空間を自在に有効利用することができなかつた。

【0009】 ここで、上記文献1では、OSが管理する自己の装置内の複数ディスクを論理ボリュームとして管理するものであり、他の装置のディスクを使用してはいない。

【0010】 また、文献2では、ネットワークを介して他のディスクを利用しているものの、当該装置のOS自身のレベルでその制御を行っていたため、複数のディスクを統合して管理するという論理ボリュームの利点を有効活用できなかつた。従って、例えば、論理ボリュームが構築されているシステムが動作不能の障害に陥った場合に、全てのディスクにアクセス不能となってしまうという問題があつた。つまり、OSレベルとディスクI/Oレベルとの間に論理ボリュームを存在させることが要請されていた。

【0011】 本発明は、上記従来の問題点に鑑みなされたものであり、マルチコンピュータ環境上の空きディスクスペースの有効利用を図ることのできる論理ボリュームシステムを提供することを目的とする。

【0012】

【課題を解決するための手段】 上記目的を達成するため

に、請求項1記載の発明は、オペレーティングシステム部と、少なくとも1つのディスク装置と、そのディスク装置を管理するディスクI/O管理部と、ネットワークバスに接続されたネットワークI/O管理部と、を含むコンピュータ装置が、前記ネットワークバスに複数接続されたシステムにおいて、少なくとも1つのコンピュータ装置において、前記オペレーティングシステム部と前記ディスクI/O管理部との間に、システム内にある1又は複数の任意のディスク装置の記憶空間を、仮想的な記憶空間である論理ボリュームとして統合管理する論理ボリューム管理部が設けられ、前記オペレーティングシステム部は、前記論理ボリュームをディスクI/O管理部と同等とみなしてディスク制御の命令を出し、前記論理ボリューム管理部は、前記オペレーティングシステム部からのディスク制御の命令に対応する現実のディスク装置を判断して当該ディスク装置の制御を行うことを特徴とする。

【0013】また、請求項2記載の発明は、上記構成において、前記論理ボリューム管理部と前記ネットワークI/O管理部とが直接接続されたことを特徴とする。

【0014】

【作用】上記構成によれば、オペレーティングシステム部とディスクI/O管理部との間に論理ボリューム管理部が設けられ、オペレーティングシステム部は、論理ボリュームを自己の保有するディスクとみなして制御を行う。すなわち、論理ボリュームは、当該コンピュータ装置及び他のコンピュータ装置におけるディスク装置の記憶空間を合わせたものであり、オペレーティングシステム部は、物理的ボリューム（ディスク装置）を意識することなく、仮想上のディスク装置（論理ボリューム）に対し、ディスク制御を行う。

【0015】この場合、論理ボリューム管理部は、オペレーティングシステム部が制御しようとする対象が実際にはどのディスク装置であるかを判断し、当該ディスク装置の制御（書き込み制御、又は読み出し制御）を行う。その制御対象が当該コンピュータ装置のディスク装置であれば、ディスクI/O管理部を介して、そのディスク装置を接続して制御を実行し、他のコンピュータ装置であれば、当該コンピュータ装置のネットワークI/O管理部及びネットワークバスを介して、前記他のコンピュータ装置のディスク装置を接続して制御を実行する。

【0016】本発明では、オペレーティングシステムレベルとI/Oレベルとの間に、論理ボリュームを存在させることができるので、オペレーティングシステム部の負担を軽減し、オペレーティングシステム部の構築時の煩雑さを軽減できる。

【0017】

【実施例】図1は、本発明に係る論理ボリュームシステムの第1実施例を示すハードウェア構成の概略ブロック

図である。

【0018】このシステムは、例えばLAN（ローカル・エリア・ネットワーク）を構成するもので、論理ボリュームバス100には、複数台の独立したコンピュータ装置101, 102が接続されている。

【0019】コンピュータ装置101には、その装置全体の動作を制御するCPU101aと、RAM等のメモリ101bと、ディスク101c用のディスク制御回路101dと、ネットワークバス100とのインターフェイスを行うネットインターフェイス回路101eと、が設けられ、それらが内部バス101fに接続されている。

【0020】同様に、コンピュータ装置102には、CPU102a、メモリ102b、ディスク102c、102d用のディスク制御回路102e、及びネットインターフェイス回路102fが設けられ、それらが内部バス102gに接続されている。

【0021】このように構成されるシステムの動作を、図2及び図3を用いて説明する。なお、図2は、本実施例のソフトウェア動作時のデータフローを示す図であり、図3は、本実施例の実行フローを示すフローチャートである。また、図2中の101, 102は、図1に示すコンピュータ装置であり、103はその他のコンピュータ装置である。

【0022】図2に示すクライアント側のコンピュータ装置102において、ユーザから、データ書き込み又はデータ読み出しの命令があると、オペレーティングシステム(OS)は、バス（ソフトウェア間の実行移行ルート）201からディスクドライブインターフェイス202を通して、論理ボリューム(DLVM)管理部203へ、ディスク制御（読み出し制御、書き込み制御）を命令する（ステップS301）。ここで、この論理ボリューム管理部203には、仮想的な論理ボリューム203aが存在している。論理ボリューム管理部203は、前記OSの命令対象が論理ボリューム203aのどの部分であるかを判断し、アクセスすべきディスク装置及びその位置を判断する（ステップS302）。ここで、論理ボリューム203aには、仮想上、コンピュータ装置101のディスク装置101cの記憶内容であるデータAを含んでいる。

【0023】なお、論理ボリューム管理部203には、論理ボリュームの各アドレスと現実の各ディスク装置の記憶空間におけるアドレスとを対応付けるテーブルなどが含まれる。また、OSからの制御命令の解読部、及び、ディスク装置への制御命令出力部、などが含まれる。

【0024】このように、論理ボリューム管理部203は、アクセスすべきデバイス（ディスク装置）がリモート、つまりネットワークバス100上の他の装置であるか否かを判断し（ステップS303）、アクセス対象が

5

自己のディスク装置であった場合は、論理ボリューム管理部203が、ディスクI/Oドライバを介して、そのまま自己のディスク装置102c又は102dの制御を行い(ステップS304)、本ルーチンを終了する。

【0025】ステップS303の判断結果が他のコンピュータ装置のディスク装置であった場合には、前記制御命令に対して、ネットワークプロトコル部204でプロトコルを合わせるための、所定のデータの付加等を施し、作成された信号がネットワークI/Oドライバ205を介して、ネットワークバス100に送出される(ステップS305)。

【0026】このネットワークバス100に送出された信号は、リモートされたサーバー側の装置(本実施例ではコンピュータ装置101)に取り込まれる。すなわち、コンピュータ装置101は、ネットワークバス100からネットワークソフトウェアを通して、前記信号を受け取る。具体的には、ネットワークI/Oドライバ206を通じて、ネットワークプロトコル部207でプロトコル処理され、ディスク制御命令が取り出され、それがオペレーティングシステム(OS)を介してサーバプロセス管理部208に通知される(ステップS306)。

【0027】これを受けたサーバプロセス管理部208は、自己のディスク装置に対するI/O命令を実行すべく(ステップS307)、バス209でOSに対するディスクI/Oのシステムコールを行う。OSは、このコールによって、インターフェイス210を介し、さらに論理ボリューム管理部211を介して、ディスクI/Oドライバへディスク制御(ディスクI/O)命令を出す。

【0028】その結果、ディスク装置101cが選択されて、前記コンピュータ装置102からディスク装置101cのデータ読み出し等が実行される。

【0029】なお、論理ボリュームは、必ずしもサーバー側のコンピュータ装置に存在しなくてもよい。

【0030】すなわち、本発明は、OSレベルとI/Oレベルとの間に論理ボリュームを存在させ、ネットワーク上のディスク記憶空間をOSからみて簡易に有効活用するものである。

【0031】図4は、第2実施例の構成を示す図であり、図2と共に要素には同一の符号が付されている。

【0032】コンピュータ装置102の論理ボリューム管理部103に仮想的に存在する論理ボリューム203bには、データA及びデータBが登録されている。ここで、データAは、ディスク装置102cとディスク装置103aとに現実に記憶されたデータA'、A''と同一のものである。また、これと同様に、論理ボリューム203bに登録されたデータは、ディスク装置102d及び101cに現実に記憶されたデータB'、B''と同一のものである。つまり、この実施例では、データを二重

10

20

30

40

50

6

化して記憶している。この場合において、論理ボリューム管理部203は、あるデータの読み出し命令を受け取ったとき、通常は、同じデータが記憶されたディスク装置のうちの一方のディスク装置を選択するが、一方のディスク装置が故障している場合には、他方のディスク装置の当該データを読み出す。なお、必要があれば、いわゆる2つのデータのブループを行って、データの信頼性を確認しても良い。

【0033】図5は、本発明の第3実施例の構成を示す図であり、図2と共に要素には同一の符号が付されている。

【0034】この実施例は、より大きな単一のファイルを分散して記憶した場合の例であり、論理ボリューム203cには、連続されたデータA-1、A-2、A-3が存在する。これに対応する物理ディスクは、自己のディスク装置102d及び他のディスク装置103b、103cであり、ディスク102dにはデータA-1、ディスク103b、103cにはデータA-2、A-3が格納されている。

【0035】図6は、本発明の第4実施例の構成を示す図であり、図2と共に要素には同一の符号が付されている。

【0036】この実施例は、ディスクストライピング方式でデータ転送の高速化を実現した例であり、論理ボリューム203dの分割された領域にはデータA～Eが存在する。これに対応する物理ディスクは、他のコンピュータ装置101、103内のディスク101c、103a、103b、103cに存在し、そして、前記第1実施例とほぼ同様の動作でディスク101cにはデータA、E、ディスク103a、103b、103cにはデータB、C、Dが前記第1実施例とほぼ同様の動作でそれぞれ分散格納されている。

【0037】

【発明の効果】以上説明したように、本発明によれば、ネットワーク上において空きディスク空間がある場合に、オペレーティング部の構造を複雑化させることなく、その空き記憶空間を有効利用できる。

【図面の簡単な説明】

【図1】本発明に係る論理ボリュームを有するネットワークシステムのハードウェア構成を示すブロック図である。

【図2】第1実施例の構成及びデータフローを示す概念図である。

【図3】第1実施例の実行フローを示すフローチャートである。

【図4】第2実施例の構成及びデータフローを示す概念図である。

【図5】第3実施例の構成及びデータフローを示す概念図である。

【図6】第4実施例の構成及びデータフローを示す概念

図である。

【符号の説明】

100 ネットワークバス

101, 102, 103 コンピュータ装置

203 論理ボリューム管理部

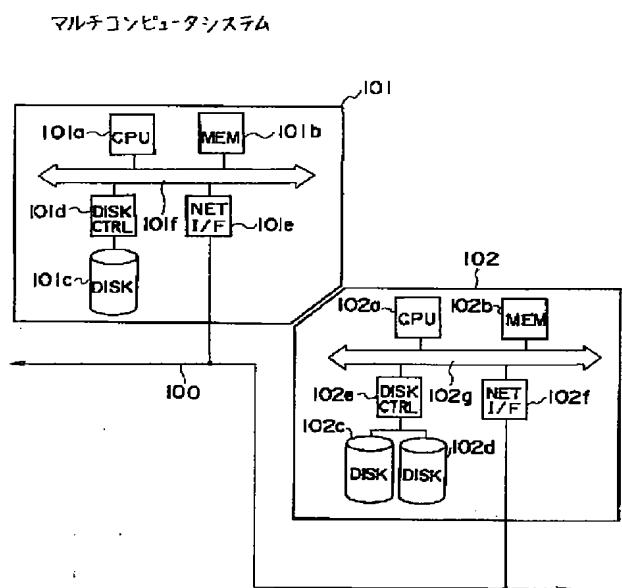
101c, 102c, 102d, 103a, 103b,

103c ディスク装置

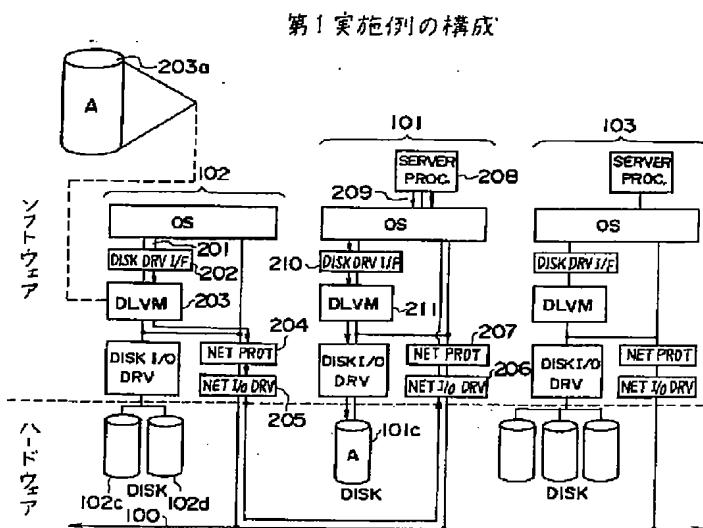
203a, 203b, 203c, 203d 論理ボリュ

ーム

【図1】

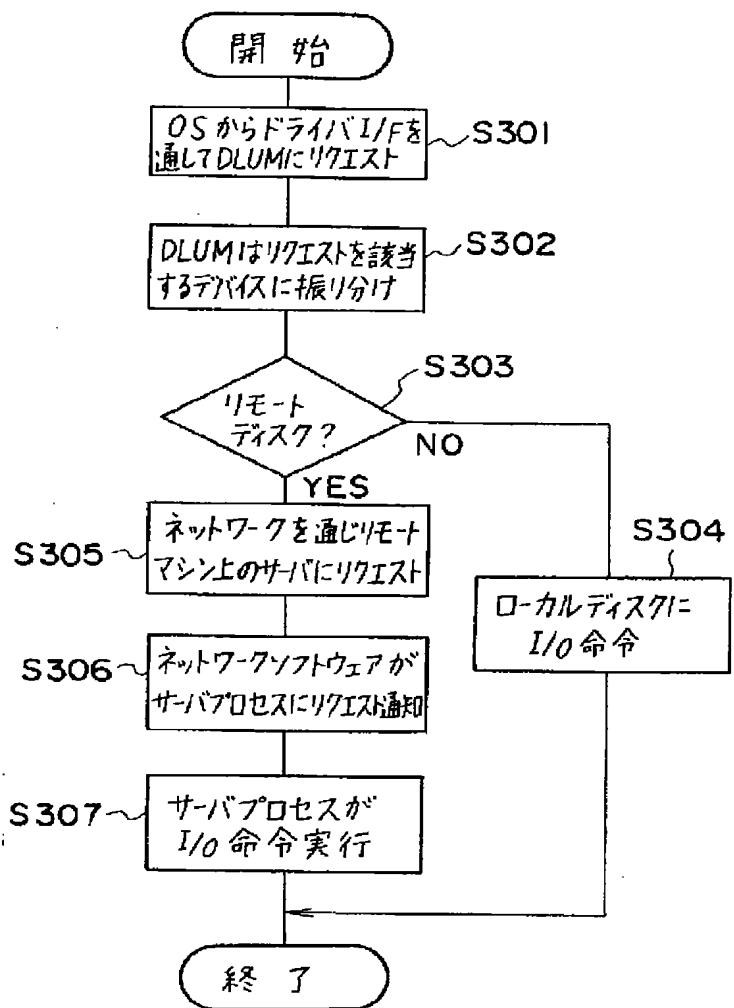


【図2】



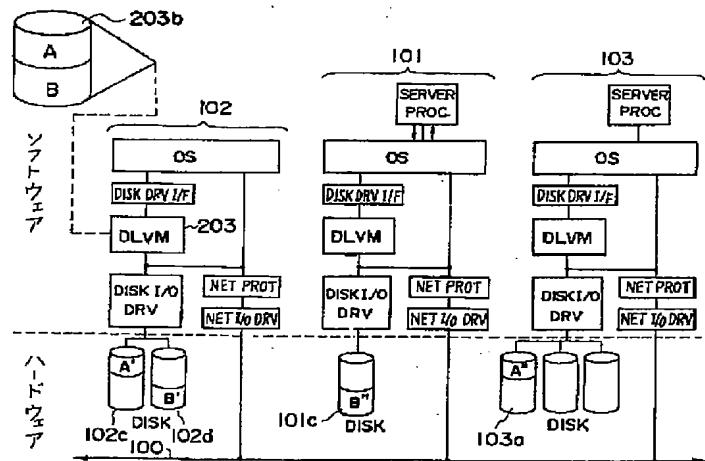
【図3】

第1実施例の実行フロー



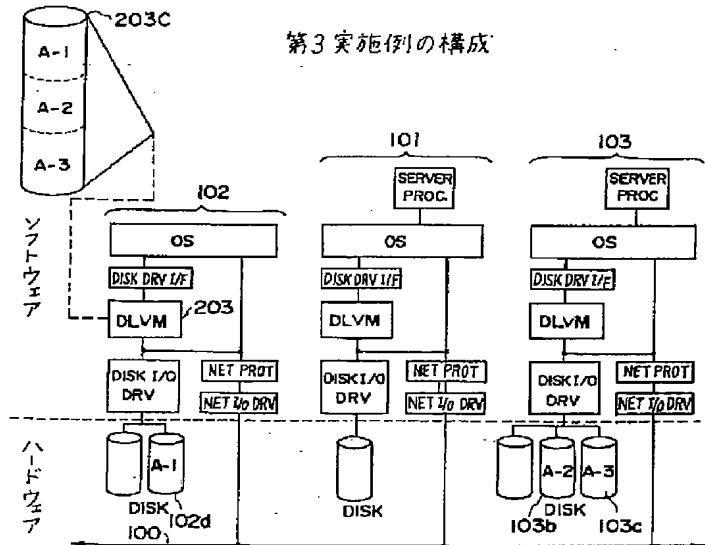
【図4】

第2実施例の構成

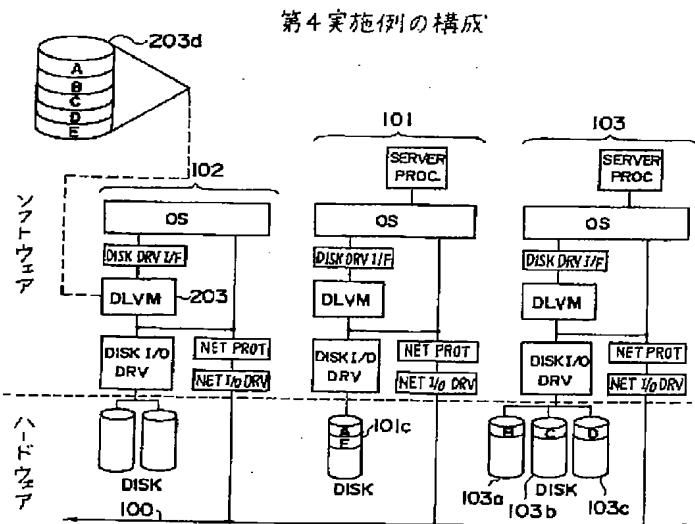


【図5】

第3 実施例の構成



【図6】



【公報種別】特許法第17条の2の規定による補正の掲載

【部門区分】第6部門第3区分

【発行日】平成9年(1997)2月14日

【公開番号】特開平5-334006

【公開日】平成5年(1993)12月17日

【年通号数】公開特許公報5-3341

【出願番号】特願平4-144622

【国際特許分類第6版】

G06F 3/06 301

【F1】

G06F 3/06 301 K 7323-5E

【手続補正書】

【提出日】平成8年3月27日

【手続補正1】

【補正対象書類名】明細書

【補正対象項目名】発明の名称

【補正方法】変更

【補正内容】

【発明の名称】論理ボリューム装置

【手続補正2】

【補正対象書類名】明細書

【補正対象項目名】特許請求の範囲

【補正方法】変更

【補正内容】

【特許請求の範囲】

【請求項1】データを記憶する記憶装置と、

ネットワークを介して接続された他計算機が有し、データを記憶する記憶装置と、上記記憶装置と、を2次記憶資源として管理し、上記他計算機の記憶装置と上記記憶装置とを用いて、論理ボリュームの管理を行なう管理手段と、

を備えたことを特徴とする論理ボリューム装置。

【請求項2】上記管理手段は、論理ボリュームの管理として、多重化、並列化、又は分割化を行なうことを特徴とする請求項1記載の論理ボリューム装置。

【手続補正3】

【補正対象書類名】明細書

【補正対象項目名】発明の詳細な説明

【補正方法】変更

【補正内容】

【発明の詳細な説明】

【0001】

【産業上の利用分野】本発明は、現実のディスク装置の記憶領域を仮想的な論理ボリュームとして管理する論理ボリューム装置に関する。

【0002】

【従来の技術】複数のコンピュータ装置がネットワークバスに接続されたシステムが各種の分野で活用されてい

る。

【0003】このようなシステムにおいて、各コンピュータ装置は、一般的に1又は複数のディスク装置を備えている。しかし、それぞれのコンピュータ装置において、各ディスク装置における記憶領域の使用率はばらつきがあるのが通常であり、記憶領域が不足するコンピュータ装置もあれば、記憶領域が余剰しているコンピュータ装置もある。

【0004】ところで、複数のディスク装置の制御を簡易にするためのものとして、従来からいわゆる「論理ボリューム」という仮想記憶領域が知られている。この論理ボリュームは、通常、現実の複数のディスク装置（物理ボリューム）の記憶領域を合わせた仮想的記憶領域であり、仮想上の大きなディスク装置に相当するものである。すなわち、コンピュータの基本システムからは、論理ボリュームだけがみえるようにし、その基本システムが個々のディスク装置を直接制御する煩雑さを解消するものである。

【0005】従来、このような論理ボリュームの適応例としては、高速化のためにディスクの分割使用を行うディスクストライピング手法、より広い連続領域を得るためにディスクの並列使用を行うディスクアレイ手法あるいはディスク障害によるファイル破損に対応するためにディスクを多重化するミラーディスク手法等が提案されている。

【0006】また、論理ボリュームに関し、特開昭63-10250号公報（文献1）や特開平3-92942号公報（文献2）等に開示されるものが挙げられる。ここで、文献1ではファイル作成時に、該当する論理ボリュームを構成する物理ボリュームを略均等に選択するようになり、ファイルを各物理ボリュームに分散することにより、論理ボリュームに対応する物理ボリュームのアクセス分散を行い、アクセス効率の向上を図っている。また、文献2では、ファイルをディスク内に任意の論理的な単位で格納、多重化するようにし、ディスクの利用効率の向上を図っている。

【0007】

【発明が解決しようとする課題】しかしながら、従来においては、マルチコンピュータ環境上で、ネットワークを介した論理ボリュームは構築されていなかった。このため、システムにおけるディスクの空き空間を自在に有効利用することができなかった。

【0008】ここで、上記文献1では、OSが管理する自己の装置内の複数ディスクを論理ボリュームとして管理するものであり、他の装置のディスクを使用してはいない。

【0009】また、文献2では、ネットワークを介して他のディスクを利用しているものの、当該装置のOS自身のレベルでその制御を行っていたため、複数のディスクを統合して管理するという論理ボリュームの利点を有効活用できなかった。従って、例えば、論理ボリュームが構築されているシステムが動作不能の障害に陥った場合に、全てのディスクにアクセス不能となってしまうという問題があった。つまり、OSレベルとディスクI/Oレベルとの間に論理ボリュームを存在させることが要請されていた。

【0010】

【0011】本発明は、上記従来の問題点に鑑みなされたものであり、論理ボリューム装置に、ネットワークを介して接続された他コンピュータ装置上のディスク装置を管理するためのアクセス機能を付加することにより、ネットワークを介しての他コンピュータ装置上のディスク装置へのミラーリングを可能にし、これにより、障害発生時にもミラーデータの利用を可能にし、より大きな論理ボリュームの管理を可能とし、あるいはネットワークを介してのディスク入出力の負荷分散等を図ることを目的とする。

【0011】

【課題を解決するための手段】上記目的を達成するために、請求項1記載の発明は、データを記憶する記憶装置と、ネットワークを介して接続された他計算機が有し、データを記憶する記憶装置と、上記記憶装置と、を2次記憶資源として管理し、上記他計算機の記憶装置と上記記憶装置とを用いて、論理ボリュームの管理を行なう管理手段と、を備えたことを特徴とする。

【0012】また、請求項2記載の発明は、論理ボリュームの管理として、多重化、並列化、又は分割化を行なうことを特徴とする。

【0013】

【作用】上記構成によれば、管理手段が当該計算機の記憶装置と他計算機の記憶装置とを併せて論理ボリュームとして管理し、ネットワークを介して論理ボリュームを構築できる。よって、ユーザーは実際のディスク装置の所在を意識することなくアクセスでき、合理的かつ効率的なデータ管理を実現できる。

【0014】また、本発明によれば、多重化を行うミラ

ーリング、並列化を行うコンカティネーティング又は分割化を行うストライピングを論理ボリュームの管理に適用できる。

【0015】なお、本発明では、オペレーティングシステムレベルとI/Oレベルとの間に、論理ボリュームを存在させることもでき、かかる構成によれば、オペレーティングシステム部の負担を軽減し、オペレーティングシステム部の構築時の煩雑さを軽減できる。

【0016】

【実施例】図1は、本発明に係る論理ボリューム装置の第1実施例を示すハードウェア構成の概略ブロック図である。

【0017】この装置は、例えばLAN(ローカル・エリア・ネットワーク)を構成するもので、論理ボリュームバス100には、複数台の独立したコンピュータ装置101, 102が接続されている。

【0018】コンピュータ装置101には、その装置全体の動作を制御するCPU101aと、RAM等のメモリ101bと、ディスク101c用のディスク制御回路101dと、ネットワークバス100とのインターフェイスを行うネットインターフェイス回路101eと、が設けられ、それらが内部バス101fに接続されている。

【0019】同様に、コンピュータ装置102には、CPU102a、メモリ102b、ディスク102c、102d用のディスク制御回路102e、及びネットインターフェイス回路102fが設けられ、それらが内部バス102gに接続されている。

【0020】このように構成される装置の動作を、図2及び図3を用いて説明する。なお、図2は、本実施例のソフトウェア動作時のデータフローを示す図であり、図3は、本実施例の実行フローを示すフローチャートである。また、図2中の101, 102は、図1に示すコンピュータ装置であり、103はその他のコンピュータ装置である。

【0021】図2に示すクライアント側のコンピュータ装置102において、ユーザから、データ書き込み又はデータ読み出しの命令があると、オペレーティングシステム(OS)は、バス(ソフトウェア間の実行移行ルート)201からディスクドライバインターフェイス202を通して、論理ボリューム(DLVM)管理部203へ、ディスク制御(読み出し制御、書き込み制御)を要求する(ステップS301)。ここで、この論理ボリューム管理部203には、仮想的な論理ボリューム203aが存在している。論理ボリューム管理部203は、前記OSの命令対象が論理ボリューム203aのどの部分であるかを判断し、アクセスすべきディスク装置及びその位置を判断する(ステップS302)。ここで、論理ボリューム203aは、コンピュータ装置101のディスク装置101cの記憶内容であるデータAを含んでい

る。

【0022】なお、論理ボリューム管理部203には、論理ボリュームの各アドレスと現実の各ディスク装置の記憶空間におけるアドレスとを対応付けるテーブルなどが含まれる。また、OSからの制御命令の解読部、及び、ディスク装置への制御命令出力部、などが含まれる。

【0023】このように、論理ボリューム管理部203は、アクセスすべきデバイス（ディスク装置）がリモート、つまりネットワークバス100上の他の装置であるか否かを判断し（ステップS303）、アクセス対象が自己のディスク装置であった場合は、論理ボリューム管理部203が、ディスクI/Oドライバを介して、そのまま自己のディスク装置102c又は102dの制御を行い（ステップS304）、本ルーチンを終了する。

【0024】ステップS303の判断結果が他のコンピュータ装置のディスク装置であった場合には、前記制御命令に対して、ネットワークプロトコル部204でプロトコルを合わせるための、所定のデータの付加等を施し、作成された信号がネットワークI/Oドライバ205を介して、ネットワークバス100に送出される（ステップS305）。

【0025】このネットワークバス100に送出された信号は、コンピュータ装置101に取り込まれる。すなわち、コンピュータ装置101は、ネットワークバス100からネットワークソフトウェアを通して、前記信号を受け取る。具体的には、ネットワークI/Oドライバ206を通じて、ネットワークプロトコル部207でプロトコル処理され、ディスク制御命令が取り出され、それがオペレーティングシステム（OS）を介してサーバプロセス管理部208に通知される（ステップS306）。

【0026】これを受けたサーバプロセス管理部208は、自己のディスク装置に対するI/O命令を実行すべく（ステップS307）、バス209でOSに対するディスクI/Oのシステムコールを行う。OSは、このコールによって、インターフェイス210を介し、さらに論理ボリューム管理部211を介して、ディスク1I/Oドライバへディスク制御（ディスクI/O）命令を出す。

【0027】その結果、ディスク装置101cが選択されて、前記コンピュータ装置102からディスク装置101cのデータ読み出し等が実行される。

【0028】なお、論理ボリュームは、必ずしもサーバ一側のコンピュータ装置に存在しなくともよい。

【0029】すなわち、本実施例によればユーザはアクセスする論理ボリュームを構成するディスク装置が、自己のディスク装置であるか、リモートのディスク装置であるかを意識せず、その実体がリモートにある場合においても、自己のディスク装置と同様にアクセスすること

が可能である。

【0030】図4は、第2実施例の構成を示す図であり、図2と共通の要素には同一の符号が付されている。

【0031】コンピュータ装置102の論理ボリューム管理部103に仮想的に存在する論理ボリューム203bには、データA及びデータBが登録されている。ここで、データAは、ディスク装置102cとディスク装置103aとに現実に記憶されたデータA'、A''と同一のものである。また、これと同様に、論理ボリューム203bに登録されたデータは、ディスク装置102d及び101cに現実に記憶されたデータB'、B''と同一のものである。つまり、この実施例では、データを二重化して記憶している。この場合において、論理ボリューム管理部203は、あるデータの読み出し命令を受け取ったとき、通常は、同じデータが記憶されたディスク装置のうちの一方のディスク装置を選択するが、一方のディスク装置が故障している場合には、他方のディスク装置の該当データを読み出す。また、上述から明らかなように、書き込みの場合は、二重化されている全てのディスク装置の該当する領域へ同一データの書き込みが行われる。本実施例によって、二重化データをリモートの他のコンピュータ装置に配置可能にすることにより、二重化ディスク装置を保有するいづれかのコンピュータ装置が故障した場合も、稼働しているコンピュータ装置を利用して該ディスクデータにアクセスすることが可能になる。さらに、本実施例の適用は二重化に限るものではなく、さらなる多重化も可能である。

【0032】図5は、本発明の第3実施例の構成を示す図であり、図2と共通の要素には同一の符号が付されている。

【0033】この実施例は、より大きな単一のファイルを分散して記憶した場合の例であり、論理ボリューム203cには、連続されたデータA-1、A-2、A-3が存在する。これに対応する物理ディスクは、自己のディスク装置102d及び他のディスク装置103b、103cであり、ディスク102dにはデータA-1、ディスク103b、103cにはデータA-2、A-3が格納されている。このように、本実施例によれば、上述から明らかなようにリモートのディスク装置をも並列使用して、コンカティネーティッド論理ボリュームを構成することにより、より大きな論理ボリュームを実現することができる。

【0034】図6は、本発明の第4実施例の構成を示す図であり、図2と共通の要素には同一の符号が付されている。

【0035】この実施例は、ディスクストライピング方式でデータ転送の高速化を実現した例であり、論理ボリューム203dの分割された領域にはデータA～Eが存在する。これに対応する物理ディスクは、他のコンピュータ装置101、103内のディスク101c、103

a, 103b, 103cに存在し、そして、前記第1実施例とほぼ同様の動作でディスク101cにはデータA, E、ディスク103a, 103b, 103cにはデータB, C, Dがそれぞれ分散格納されている。

【0036】このように、本実施例によれば、ストライプ構成（分割化）された論理ボリュームにおいて、各ストライプに割り付けられたディスク装置をネットワークを介して並列に動作させることにより、より高い、論理ボリュームにおける入出力性能を実現することができる。

【0037】

【発明の効果】以上説明したように、本発明によれば、論理ボリューム装置に、ネットワークを介して接続された他コンピュータ装置上のディスク装置を管理するためのアクセス機能を付加したので、コンピュータ装置故障時にもディスクアクセスを可能にでき、またより大きな論理ボリュームを管理でき、さらにネットワークを介してのディスク入出力の負荷分散化を図れるという効果がある。

【公報種別】特許法第17条の2の規定による補正の掲載

【部門区分】第6部門第3区分

【発行日】平成10年(1998)11月13日

【公開番号】特開平5-334006

【公開日】平成5年(1993)12月17日

【年通号数】公開特許公報5-3341

【出願番号】特願平4-144622

【国際特許分類第6版】

G06F 3/06 301

【F I】

G06F 3/06 301 K

【手続補正書】

【提出日】平成8年3月27日

【手続補正1】

【補正対象書類名】明細書

【補正対象項目名】発明の名称

【補正方法】変更

【補正内容】

【発明の名称】論理ボリューム装置

【手続補正2】

【補正対象書類名】明細書

【補正対象項目名】特許請求の範囲

【補正方法】変更

【補正内容】

【特許請求の範囲】

【請求項1】データを記憶する記憶装置と、

ネットワークを介して接続された他計算機が有し、データを記憶する記憶装置と、上記記憶装置と、を2次記憶資源として管理し、上記他計算機の記憶装置と上記記憶装置とを用いて、論理ボリュームの管理を行なう管理手段と、

を備えたことを特徴とする論理ボリューム装置。

【請求項2】上記管理手段は、論理ボリュームの管理として、多重化、並列化、又は分割化を行なうことを特徴とする請求項1記載の論理ボリューム装置。

【手続補正3】

【補正対象書類名】明細書

【補正対象項目名】発明の詳細な説明

【補正方法】変更

【補正内容】

【発明の詳細な説明】

【0001】

【産業上の利用分野】本発明は、現実のディスク装置の記憶領域を仮想的な論理ボリュームとして管理する論理ボリューム装置に関する。

【0002】

【従来の技術】複数のコンピュータ装置がネットワークバスに接続されたシステムが各種の分野で活用されてい

る。

【0003】このようなシステムにおいて、各コンピュータ装置は、一般的に1又は複数のディスク装置を備えている。しかし、それぞれのコンピュータ装置において、各ディスク装置における記憶領域の使用率はばらつきがあるのが通常であり、記憶領域が不足するコンピュータ装置もあれば、記憶領域が余剰しているコンピュータ装置もある。

【0004】ところで、複数のディスク装置の制御を簡易にするためのものとして、従来からいわゆる「論理ボリューム」という仮想記憶領域が知られている。この論理ボリュームは、通常、現実の複数のディスク装置（物理ボリューム）の記憶領域を合わせた仮想的記憶領域であり、仮想上の大きなディスク装置に相当するものである。すなわち、コンピュータの基本システムからは、論理ボリュームだけがみえるようにし、その基本システムが個々のディスク装置を直接制御する煩雑さを解消するものである。

【0005】従来、このような論理ボリュームの適応例としては、高速化のためにディスクの分割使用を行うディスクストライピング手法、より広い連続領域を得るためにディスクの並列使用を行うディスクアレイ手法あるいはディスク障害によるファイル破損に対応するためにディスクを多重化するミラーディスク手法等が提案されている。

【0006】また、論理ボリュームに関し、特開昭63-10250号公報（文献1）や特開平3-92942号公報（文献2）等に開示されるものが挙げられる。ここで、文献1ではファイル作成時に、該当する論理ボリュームを構成する物理ボリュームを略均等に選択するようになり、ファイルを各物理ボリュームに分散することにより、論理ボリュームに対応する物理ボリュームのアクセス分散を行い、アクセス効率の向上を図っている。また、文献2では、ファイルをディスク内に任意の論理的な単位で格納、多重化するようにし、ディスクの利用効率の向上を図っている。

【0007】

【発明が解決しようとする課題】しかしながら、従来においては、マルチコンピュータ環境上で、ネットワークを介した論理ボリュームは構築されていなかった。このため、システムにおけるディスクの空き空間を自在に有効利用することができなかつた。

【0008】ここで、上記文献1では、OSが管理する自己の装置内の複数ディスクを論理ボリュームとして管理するものであり、他の装置のディスクを使用してはいなない。

【0009】また、文献2では、ネットワークを介して他のディスクを利用しているものの、当該装置のOS自体のレベルでその制御を行っていたため、複数のディスクを統合して管理するという論理ボリュームの利点を有効活用できなかつた。従って、例えば、論理ボリュームが構築されているシステムが動作不能の障害に陥った場合に、全てのディスクにアクセス不能となってしまうという問題があつた。つまり、OSレベルとディスクI/Oレベルとの間に論理ボリュームを存在させることが要請されていた。

【0010】

【0011】本発明は、上記従来の問題点に鑑みなされたものであり、論理ボリューム装置に、ネットワークを介して接続された他コンピュータ装置上のディスク装置を管理するためのアクセス機能を付加することにより、ネットワークを介しての他コンピュータ装置上のディスク装置へのミラーリングを可能にし、これにより、障害発生時にもミラーデータの利用を可能にし、より大きな論理ボリュームの管理を可能とし、あるいはネットワークを介してのディスク入出力の負荷分散等を図ることを目的とする。

【0011】

【課題を解決するための手段】上記目的を達成するために、請求項1記載の発明は、データを記憶する記憶装置と、ネットワークを介して接続された他計算機が有し、データを記憶する記憶装置と、上記記憶装置と、を2次記憶資源として管理し、上記他計算機の記憶装置と上記記憶装置とを用いて、論理ボリュームの管理を行なう管理手段と、を備えたことを特徴とする。

【0012】また、請求項2記載の発明は、論理ボリュームの管理として、多重化、並列化、又は分割化を行なうことを特徴とする。

【0013】

【作用】上記構成によれば、管理手段が当該計算機の記憶装置と他計算機の記憶装置とを併せて論理ボリュームとして管理し、ネットワークを介して論理ボリュームを構築できる。よつて、ユーザーは実際のディスク装置の所在を意識することなくアクセスでき、合理的かつ効率的なデータ管理を実現できる。

【0014】また、本発明によれば、多重化を行うミラ

ーリング、並列化を行うコンカティネーティング又は分割化を行うストライピングを論理ボリュームの管理に適用できる。

【0015】なお、本発明では、オペレーティングシステムレベルとI/Oレベルとの間に、論理ボリュームを存在させることもでき、かかる構成によれば、オペレーティングシステム部の負担を軽減し、オペレーティングシステム部の構築時の煩雑さを軽減できる。

【0016】

【実施例】図1は、本発明に係る論理ボリューム装置の第1実施例を示すハードウェア構成の概略ブロック図である。

【0017】この装置は、例えばLAN(ローカル・エリア・ネットワーク)を構成するもので、論理ボリュームバス100には、複数台の独立したコンピュータ装置101, 102が接続されている。

【0018】コンピュータ装置101には、その装置全体の動作を制御するCPU101aと、RAM等のメモリ101bと、ディスク101c用のディスク制御回路101dと、ネットワークバス100とのインターフェイスを行うネットインターフェイス回路101eと、が設けられ、それらが内部バス101fに接続されている。

【0019】同様に、コンピュータ装置102には、CPU102a、メモリ102b、ディスク102c、102d用のディスク制御回路102e、及びネットインターフェイス回路102fが設けられ、それらが内部バス102gに接続されている。

【0020】このように構成される装置の動作を、図2及び図3を用いて説明する。なお、図2は、本実施例のソフトウェア動作時のデータフローを示す図であり、図3は、本実施例の実行フローを示すフローチャートである。また、図2中の101, 102は、図1に示すコンピュータ装置であり、103はその他のコンピュータ装置である。

【0021】図2に示すクライアント側のコンピュータ装置102において、ユーザから、データ書き込み又はデータ読み出しの命令があると、オペレーティングシステム(OS)は、パス(ソフトウェア間の実行移行ルート)201からディスクドライバインターフェイス202を通して、論理ボリューム(DLVM)管理部203へ、ディスク制御(読み出し制御、書き込み制御)を要求する(ステップS301)。ここで、この論理ボリューム管理部203には、仮想的な論理ボリューム203aが存在している。論理ボリューム管理部203は、前記OSの命令対象が論理ボリューム203aのどの部分であるかを判断し、アクセスすべきディスク装置及びその位置を判断する(ステップS302)。ここで、論理ボリューム203aは、コンピュータ装置101のディスク装置101cの記憶内容であるデータAを含んでい

る。

【0022】なお、論理ボリューム管理部203には、論理ボリュームの各アドレスと現実の各ディスク装置の記憶空間におけるアドレスとを対応付けるテーブルなどが含まれる。また、OSからの制御命令の解読部、及び、ディスク装置への制御命令出力部、などが含まれる。

【0023】このように、論理ボリューム管理部203は、アクセスすべきデバイス（ディスク装置）がリモート、つまりネットワークバス100上の他の装置であるか否かを判断し（ステップS303）、アクセス対象が自己のディスク装置であった場合は、論理ボリューム管理部203が、ディスクI/Oドライバを介して、そのまま自己のディスク装置102c又は102dの制御を行い（ステップS304）、本ルーチンを終了する。

【0024】ステップS303の判断結果が他のコンピュータ装置のディスク装置であった場合には、前記制御命令に対して、ネットワークプロトコル部204でプロトコルを合わせるための、所定のデータの付加等を施し、作成された信号がネットワークI/Oドライバ205を介して、ネットワークバス100に送出される（ステップS305）。

【0025】このネットワークバス100に送出された信号は、コンピュータ装置101に取り込まれる。すなわち、コンピュータ装置101は、ネットワークバス100からネットワークソフトウェアを通して、前記信号を受け取る。具体的には、ネットワークI/Oドライバ206を通じて、ネットワークプロトコル部207でプロトコル処理され、ディスク制御命令が取り出され、それがオペレーティングシステム（OS）を介してサーバプロセス管理部208に通知される（ステップS306）。

【0026】これを受けたサーバプロセス管理部208は、自己のディスク装置に対するI/O命令を実行すべく（ステップS307）、バス209でOSに対するディスクI/Oのシステムコールを行う。OSは、このコールによって、インターフェイス210を介し、さらに論理ボリューム管理部211を介して、ディスクI/Oドライバへディスク制御（ディスクI/O）命令を出す。

【0027】その結果、ディスク装置101cが選択されて、前記コンピュータ装置102からディスク装置101cのデータ読み出し等が実行される。

【0028】なお、論理ボリュームは、必ずしもサーバ一側のコンピュータ装置に存在しなくてもよい。

【0029】すなわち、本実施例によればユーザはアクセスする論理ボリュームを構成するディスク装置が、自己のディスク装置であるか、リモートのディスク装置であるかを意識せず、その実体がリモートにある場合においても、自己のディスク装置と同様にアクセスすること

が可能である。

【0030】図4は、第2実施例の構成を示す図であり、図2と共に要素には同一の符号が付されている。

【0031】コンピュータ装置102の論理ボリューム管理部103に仮想的に存在する論理ボリューム203bには、データA及びデータBが登録されている。ここで、データAは、ディスク装置102cとディスク装置103aとに現実に記憶されたデータA'、A''と同一のものである。また、これと同様に、論理ボリューム203bに登録されたデータは、ディスク装置102d及び101cに現実に記憶されたデータB'、B''と同一のものである。つまり、この実施例では、データを二重化して記憶している。この場合において、論理ボリューム管理部203は、あるデータの読み出し命令を受け取ったとき、通常は、同じデータが記憶されたディスク装置のうちの一方のディスク装置を選択するが、一方のディスク装置が故障している場合には、他方のディスク装置の当該データを読み出す。また、上述から明らかなように、書き込みの場合は、二重化されている全てのディスク装置の該当する領域へ同一データの書き込みが行われる。本実施例によって、二重化データをリモートの他のコンピュータ装置に配置可能にすることにより、二重化ディスク装置を保有するいざれかのコンピュータ装置が故障した場合も、稼働しているコンピュータ装置を利用して該ディスクデータにアクセスすることが可能になる。さらに、本実施例の適用は二重化に限るものではなく、さらなる多重化も可能である。

【0032】図5は、本発明の第3実施例の構成を示す図であり、図2と共に要素には同一の符号が付されている。

【0033】この実施例は、より大きな単一のファイルを分散して記憶した場合の例であり、論理ボリューム203cには、連続されたデータA-1、A-2、A-3が存在する。これに対応する物理ディスクは、自己のディスク装置102d及び他のディスク装置103b、103cであり、ディスク102dにはデータA-1、ディスク103b、103cにはデータA-2、A-3が格納されている。このように、本実施例によれば、上述から明らかなようにリモートのディスク装置をも並列使用して、コンカティネーティッド論理ボリュームを構成することにより、より大きな論理ボリュームを実現することができる。

【0034】図6は、本発明の第4実施例の構成を示す図であり、図2と共に要素には同一の符号が付されている。

【0035】この実施例は、ディスクストライピング方式でデータ転送の高速化を実現した例であり、論理ボリューム203dの分割された領域にはデータA～Eが存在する。これに対応する物理ディスクは、他のコンピュータ装置101、103内のディスク101c、103

a, 103b, 103cに存在し、そして、前記第1実施例とほぼ同様の動作でディスク101cにはデータA, E、ディスク103a, 103b, 103cにはデータB, C, Dがそれぞれ分散格納されている。

【0036】このように、本実施例によれば、ストライプ構成（分割化）された論理ボリュームにおいて、各ストライプに割り付けられたディスク装置をネットワークを介して並列に動作させることにより、より高い、論理ボリュームにおける入出力性能を実現することができる。

【0037】

【発明の効果】以上説明したように、本発明によれば、論理ボリューム装置に、ネットワークを介して接続された他コンピュータ装置上のディスク装置を管理するためのアクセス機能を付加したので、コンピュータ装置故障時にもディスクアクセスを可能にでき、またより大きな論理ボリュームを管理でき、さらにネットワークを介してのディスク入出力の負荷分散化を図れるという効果がある。